PCT/GB 98/02628
21 SEPTEMBER 1998

The Patent Office

The Patent Office
Concept House
Cardiff Road
Newport
South Wales
NP9 1RH

09/486676

PRIORITY DOCUMENT

REC'D 28 SEP 1998
WIPO PCT

I, the undersigned, being an officer duly authorised in accordance with Section 74(1) and (4) of the Deregulation & Contracting Out Act 1994, to sign and issue certificates on behalf of the Comptroller-General, hereby certify that annexed hereto is a true copy of the documents as originally filed in connection with the patent application identified therein.

In accordance with the Patents (Companies Re-registration) Rules 1982, if a company named in this certificate and any accompanying documents has re-registered under the Companies Act 1980 with the same name as that with which it was registered immediately before re-registration save for the substitution as, or inclusion as, the last part of the name of the words "public limited company" or their equivalents in Welsh, references to the name of the company in this certificate and any accompanying documents shall be treated as references to the name with which it is so re-registered.

In accordance with the rules, the words "public limited company" may be replaced by p.l.c., plc, P.L.C. or PLC.

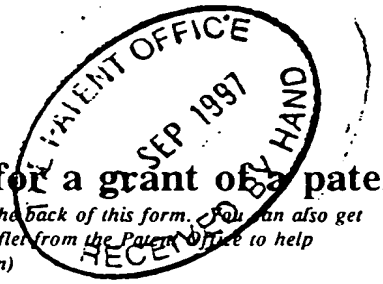
Re-registration under the Companies Act does not constitute a new legal entity but merely subjects the company to certain additional company law rules.

Signed

Dated 14 September 1998

'ents Form 1/77

Patents Act 1977
(Rule 16)



The Patent Office

& CO

03SEP97 E300073-6 002246
P01/7700 25.00 - 9718609.2

Request for a grant of a patent

(See the notes on the back of this form. You can also get an explanatory leaflet from the Patent Office to help you fill in this form)

The Patent Office

Cardiff Road
Newport
Gwent NP9 1RH

1. Your reference — P/3037.GB CLM

2. Patent application number *(The Patent Office will fill in this part)* — 9718609.2

3. Full name, address and postcode of the or of each applicant *(underline all surnames)*

Imperial College of Science, Technology & Medicine,
Sherfield Building,
Exhibition Road,
London, SW7 2AZ

Patents ADP number *(if you know it)*

If the applicant is a corporate body, give the country/state of its incorporation — United Kingdom

4. Title of the invention — Fusion Protein

5. Name of your agent *(if you have one)* — D YOUNG & CO

"Address for service" in the United Kingdom to which all correspondence should be sent *(including the postcode)*

21 NEW FETTER LANE
LONDON
EC4A 1DA

Patents ADP number *(if you have one)* — 59006

6. If you are declaring priority from one or more earlier patent applications, give the country and date of filing of the or each of these earlier applications and (if you know it) the or each application number

Country	Priority application number *(if you know it)*	Date of filing *(day/month/year)*

7. If this application is divided or otherwise derived from an earlier UK application, give the number and filing date of the earlier application

Number of earlier application	Date of filing *(day/month/year)*

8. **Is a statement of inventorship and of right to grant of a patent required in support of this request?** *(Answer 'Yes' if:*
a) any applicant named in part 3 is not an inventor, or
b) there is an inventor who is not named as an applicant, or
c) any named applicant is a corporate body.
See note (d))

Yes

9. **Enter the number of sheets for any of the following items you are filing with this form. Do not count copies of the same document**

Continuation sheets of this form	No
Description	15
Claims(s)	3
Abstract	1
Drawing(s)	1 + 1

10. **If you are also filing any of the following, state how many against each item.**

Priority documents	
Translations of priority documents	
Statement of inventorship and right to grant of a patent *(Patents Form 7/77)*	
Request for preliminary examination and search *(Patents Form 9/77)*	
Request for substantive examination *(Patents Form 10/77)*	
Any other documents *(please specify)*	

11. **I/We request the grant of a patent on the basis of this application.**

Signature D. Young & Co.

D YOUNG & CO
Agents for the Applicants

Date 2 September 1997

12. **Name and daytime telephone number of the person to contact in the United Kingdom** MALLALIEU, Catherine Louise 0171 353 434.

Warning

After an application for a patent has been filed, the Comptroller of the Patent Office will consider whether publication or communication of the invention should be prohibited or restricted under Section 22 of the Patents Act 1977. You will be informed if it is necessary to prohibit or restrict your invention in this way. Furthermore, if you live in the United Kingdom, Section 23 of the Patents Act 1977 stops you from applying for a patent abroad without first getting written permission from the Patent Office unless an application has been filed at least 6 weeks beforehand in the United Kingdom for a patent for the same invention and either no direction prohibiting publication or communication has been given, or any such direction has been revoked.

Notes

a) If you need help to fill in this form or you have any questions, please contact the Patent Office on 0645 500505

b) Write your answers in capital letters using black ink or you may type them

c) If there is not enough space for all the relevant details on any part of this form, please continue on a separte sheet of paper and write "see continuation sheet" in the relevant part(s). Any continuation sheet should be attached to this form.

d) If you answered 'Yes' Patents Form 7/77 will need to be filed.

e) Once you have filled in the form you must remember to sign and date it.

f) For details of the fee and ways to pay please contact the Patent Office.

Fusion Protein

The present invention relates to novel fusion proteins.

The gene antennapedia (*Antp*) encodes a transcriptional factor that has been shown to control antero-posterior morphogenesis in *Drosophila* embryo. The protein sequence of antennapedia is characterised by the presence of a 60 amino acids motif (homeodomain) that binds to specific DNA target elements. Antennapedia homologues have been found in nearly all multicellular organisms and show a very high degree of amino acid sequence identity. The human and drosophila antennapedia proteins differ in the sequence of the homeodomain only for one conservative amino acid substitution.

It has been observed that antennapedia and its homeodomain are able to translocate across the cytoplasmic membrane of mammalian cells. The translocation does not depend on cell endocytosis and it has been reported that translocation occurs at both 4°C and 37° C. Homeodomain synthetic peptides made of D amino acids are also able to cross the cytoplasmic membrane. This finding would rule out the possibility that *Antp* is translocated through a receptor mediated mechanism. This property has been exploited to vehiculate small viral sequences into the cytoplasm of cultured cells as well as to elicit an MHC class I restricted cytotoxic immune response against the nucleoprotein of the influenza virus. However, to date the homeodomain of *Antp* has been mainly used to transport small synthetic peptides. Schutze-Redelmeier M-P et al (1996) Introduction of Exogenous Antigens into the MHC Class I Processing and Presentation Pathway by *Drosophila* Antennapedia Homeodomain Primes Cytotoxic T Cells In Vivo. *The Journal of Immunology* 650-655 mentions that the homeodomain of the antennapedia molecule can be used to deliever up to 50 additional amino acids to the cytoplasm. Using 16 amino acids of the third helix they mention that fusion peptides containing up to 94 amino acids

could be prepared, but this is not demonstrated. Further the thrust of the disclosure is only towards delivery of synthetic peptides.

We have now found that the homeodomain of *Antp* can be used to translocate long polypeptide chains or functional or regulatory proteins. Thus we have found that *Antp* may be used to manipulate the cytoplasm of living cells by introducing, for example, functional proteins. This will have wide applicability, particualrly in the medical field.

Thus according to one aspect of the present invention there is provided a fusion protein comprising:

(a) a first region comprising the homeodomain of antennapedia or a variant thereof; and

(b) a second region not naturally associated with the first region comprising a protein of at least 100 amino acids.

Preferably the first region is at the amino terminal end of the second region.

According to another aspect of the present invention there is provided nucleic acid encoding the fusion protein of the present invention.

According to another aspect of the present invention there is provided an expression vector comprising the nucleic acid of the present invention operably linked to a promoter.

According to yet another aspect of the present invention there is provided a host cell transformed with the expression vector of the present invention.

The development of an appropriate procedure for obtaining the fusion proteins from bacterial lysate was important. With conventional methods we found that small changes

in pH, differences in osmolarity and exposure to denaturing reagents dramatically affect the translocation property of the *Antp* homeodomain.

We have now found a general method for preparing the fusion protein of the present ivnention, and in particular an appropriate procedure for purifying the fusion proteins from the bacterial lysate. This method allows fusion proteins to be obtained which will translocate a protein of at least 100 amino acids across the cell membrane.

Thus according to another aspect of the present invention there is provided a method for preparing a fusion protein comprising:

(i) culturing the host cell according to the present invention under conditions which provide for the expression of the fusion protein from the expression vector within the host cell; and

(ii) recovering the fusion protein, which method comprises fusing an amino acid tail to the fusion protein, which tail is capable of binding to at least one substrate and not to another substrate, and wherein the fusion protein is caused to bind via the tail to at least one substrate such that components of the host cell do not bind to this substrate; and the fusion protein is contacted with the other substrate such that the fusion protein is not bound and remaining components of the host cell are bound to the other substrate.

According to yet another aspect of the present invention there is provided a method for preparing a fusion protein comprising:

(i) culturing a host cell, transformed with an expression vector comprising nucleic acid, operably linked to a promoter, encoding (a) a first region comprising the homeodomain of antennapedia or a mutant thereof; and (b) a second region not naturally associated with the first region comprising an amino acid sequence of at least 100 amino acids in length, under conditions which provide for expression of the fusion protein from the expression vector within the host cell; and

(ii) recovering the fusion protein, which method comprises comprises fusing an amino acid tail to the fusion protein, which tail is capable of binding to at least one substrate and not to another substrate, and wherein the fusion protein is caused to bind via the tail to at least one substrate such that components of the host cell do not bind to this substrate; and the fusion protein is contacted with the other substrate such that the fusion protein is not bound and remaining components of the host cell are bound to the other substrate.

According to a further aspect of the present invention there is provided a method of purifying a fusion protein comprising fusing an amino acid tail to the fusion protein, which tail is capable of binding to at least one substrate and not to another substrate, and wherein the fusion protein is caused to bind via the tail to at least one substrate such that impurities do not bind to this substrate; and the fusion protein is contacted with the other substrate such that the fusion protein is not bound and remaining impurities are bound to the other substrate.

Thus in broad terms the present invention can be seen as the use of a tail attached to the fusion protein which allows both positive and negative purification steps.

Preferably the amino acid tail is fused to the carboxy terminal end of the fusion protein.

Preferably the amino acid tail comprises HHHHHHGS.

Preferably the substrate is a nickel column or an antibody with affinity for the amino acid tail.

According to a preferred embodiment the fusion protein is consecutively contacted with two substrates with which it has affinity via the amino acid tail.

Thus preferably both the nickel column and antibody are used. They may be used in any order.

Further aspects of the present invention include a fusion protein prepared by the method of the present invention; a pharmaceutical composition comprising the fusion protein of the present invention and the use of the fusion protein of the present invention in the preparation of a medicament for the treatment of a disease or infection.

Various preferred features and embodiments of the present invention will now be described by way of non-limiting example, and with reference to the accompanying drawings in which:

Figure 1 shows the structure and sequence of the antennapedia homeodomain obtainable from *Drosophila*; and

Figure 2 further shows two mutants, designated pAntp 50H and pAntp 40P2.

First Region

The first region of the fusion protein of the invention may comprise a natural or synthetic homeodomain of antennapedia.

The homeodomain of the *Antp* gene obtainable from *Drosophila* is shown in Figure 1 and in Seq ID No. 1. Sequences homologous to this homeodomain have been isolated from other organisms, including vertebrates, mammals and humans, and these are included in the present invention. The homeodomain may be prepared using standard techniques such as cloning using the procedure described in Joliet et al (1991) Antennapedia homeobox peptide regulates neural morphogenesis. *Proc. Natl. acad. Sci.* **88** 1864-1868. As previously indicated differences in the sequences of such multicellular organisms are generally conservative in nature. However, this may not necessarily be the case and other

such sequences are included in the present invention, and for example where the sequence homology is about 50% or more, e.g. 60%, 70%, 80% or 90%, with the sequence obtainable from *Drosophila*. By conservative amino acid changes we mean replacing an amino acid from one of the amino acid groups, namely hydrophobic, polar, acidic or basic, with an amino acid from within the same group. An example of such a change is the replacement of valine by methionine and vice versa.

In addition variants may be used provided that they retain the ability to translocate the membrane. Such variants may be prepared using conventional techniques such as site-directed mutagenesis. Further it may be possible to direct the second region to different cell compartments by modify the homeodomain.

Some variants of the homeodomain which retain the ability to translocate the membrane have been reported in the art and these are included in the scope of the present invention, together with any which become available. For example, Gehring W (1987) Homeo Boxes in the Study of Development. *Science* **236** 1245-1252 discloses a homeodomain of 62 amino acids, i.e. with glu at position 0 and lys at position 61. Bloch-Gallego E at al (1993) Antennapedia Homeobox Peptide Enhances Growth and Branching of Embryonic Chicken Motoneurons In Vitro. *The Journal of Cell Biology* **120(2)** 485-492 discloses a mutant called pAntp40P2 that was still able to translocate through the motoneuron membrane and to reach the nucleus. In this mutant the leucine and threonine residues in positions 40 and 41 were replaced by two proline residues. Le Roux et al (1993) Neurotropic activity of the Antennapedia homeodomain depends on its specific DNA-binding properties. *Proc. Natl. Acad. Sci.* **90** 9120-9124 generated two mutants pAntp 50A and pAntp 40P2 as shown in Figure 2 which retain the ability to translocate through the neuronal membrane. Schutze-Redelmeier M-P et al (1996) *supra* disclose that a 16 amino acid C-terminal (third helix) segment has been used to address oligonucleotides and oligopeptides to the cytoplasm and nuclei of cells in culture. However, whilst not wishing to be bound by any theory it is believed that for delivering the proteins of the present invention an amino acid sequence of about 60 is preferred.

Second Region

The second region of the fusion protein according to the invention may comprise any sequence of interest which is not naturally associated with the first region. Usually this will mean that the sequence of interest will be found in nature encoded by a gene different from the gene encoding the first region. The second region may be from the same species as the first region, but are present in the fusion protein of the invention in a manner different from the natural situation, or from a different species.

The second region of the present invention is at least 100 amino acids in length. The present invention is particularly useful for longer sequences, e.g. at least 150, 200, 400 or 1000 amino acids in length. For the avoidance of doubt the term "protein" as used herein also encompasses polypeptides of the required length.

The nucleic acid of the present invention encoding the fusion proteins of the present invention may be constructed using standard recombinant DNA methodologies. Conventional expression vectors and host cells may also be used in the invention.

Applications for the fusion proteins of the present invention include:

1. Antigen delivery system for inducing for example cytoxic immunity. Bacterial and viral antigens translocated by *Antp* in the cell cytoplasm would be processed and associated with MHC class I molecules. This antigen processing and presenting pathway is known to activate specific CD8 cytoxic lymphocytes.

2. Therapy for inherited genetic diseases. *Antp* may be used to introduce functional proteins in the cytoplasm of genetically deficient cell types.

3. Cancer therapy. *Antp* may be used to transport into cancer cells molecules that are able to restore cell cycle control or induce differentiation. For example it is understood that many cancer cells would undergo apoptosis if a functional P-53 molecule is introduced into their cytoplasm. The present invention may be used to deliver such genes.

4. Develop antibacterial and antiviral measures. For example, *Antp* may be used to transport in the cytoplasm of infected cells recombinant antibodies directed against molecules that are crucial for bacterial and viral replication.

We have now generated recombinant chimeric proteins constituted by the homeodomain of antennapedia fused either to the 85A antigen of *Mycobacterium tuberculosis* or to a single chain recombinant antibody directed against the malaria antigen Pb21.

5. Use in expression systems. For example, it is desirable to express exogenous proteins in eukaryotic cells so that they get processed correctly. However, many exogenous proteins are toxic to eukaryotic cells. In manufacturing exogenous proteins it is therefore desirable to achieve temporal expression of the exogenous protein. The system may therefore be used in connection with an inducible promoter for this or any other application involving such a system. In one such system, the loxP/cre system, the eukaryotic cell contains a gene encoding the protein of interest, functionally linked to a repressor. Thus, in the normal situation, the protein is not expressed. The *Antp* homeodomain may be fused to a transducer in accordance with the invention. Application of this fusion protein to the eukaryotic cell will thus cause expression of the protein.

In particular we have found that the fusion proteins according to the present invention were able to translocate, after purification, for example in human resting lymphocytes, macrophages as well as in cell lines, such as Hella, P815 and EL4.

Expression and purification of *Antp*-fusion proteins

Whilst not wishing to be bound by any theory we believe the ability of *Antp*-fusion proteins to translocate across the cell surface membrane may be very much dependent on the conformation of the recombinant proteins. We could not observe any translocation by using either bacteria cell extracts or purified proteins exposed to small amounts of detergent (ionic and non-ionic) or denaturating agents (urea or guanidinuim). This conformation dependent property has represented a serious problem for using *Antp*-fusion proteins in both *in vitro* and *in vivo* experiments as all reports so far published use either synthetic peptides or bacterial cell extracts. We have overcome this limitation by purifying the Antp-fusion protein under native conditions in accordance with the invention.

Example 1 - We have introduced at the carboxyl-terminal end of the recombinant protein a tail of six histidine which confers to those proteins a high affinity for nickel ions. Bacteria cells expressing the recombinant proteins were lysed by repeated exposure to ultrasounds. Cell lisates were centrifuged and loaded on a nickel column (Quiagen). The protein bound to the column were eluted step wise by changing the pH of the buffer. Under these conditions the *Antp*-fusion proteins could be selectively enriched. We have estimated by gel electrophoresis that the *Antp*-fusion proteins contributed up to 75% of the material eluted from the nickel column in the pH range between 4.0 to 5.0. The *Antp* recombinant proteins were further purified by affinity chromatography using the antibody 4 D 11 directed against the histidine tail. The recombinant proteins were eluted from the column by changing the pH of the buffer. Our results indicate that fractions eluted at pH 2.8 contained the recombinant proteins nearly free of any bacterial contaminants. These protein fraction were then loaded on a polymixine column to selectively subtract LPS and related bacterial contaminants. Gel electrophoresis analysis indicated that the eluate of the polymixine column contained the recombiant protein totally devoid of bacterial

contaminants. Purified *Antp*-fusion proteins were tested for their ability to translocate across the cell surface membrane of different cell types. Our results indicate that the purified fusion proteins Antp-85A and Antp-85A.60 are able to translocate across a variety of cell lines (Hela, Hep-G2, P815 EL4) and human monocytes. The ability of *Antp* to vehiculate large molecules across the cell membrane was not affected by the particular amino-acid composition of the fusion proteins. A different Antp-fusion molecule, generated by using the malaria antigen Thrombospondin related anonymous protein (TRAP) of about 150 amino acids, was translocated across the cell membrane as efficiently as the 85A constructs.

The monoclonal antibody 4 D11

The monoclonal antibody (MAAAb) 4 D 11 was found by screening in ELISA hydridomas generated from a mouse that was immunised with a recombinant protein containing a six histidine tail at its amino-terminal end. The antibody is available from Imperial College of Science, Technology and Medicine, c/o Impel, U.K. A molecular characterisation of the epitope showed that this Mab recognises the amino acid sequence HHHHHHGS both at the amino and at the carboxyl terminal end of recombinant proteins (**Figures to be included**). The antibody has an IgG1 isotype and can be easily purified on protein A column. Our results indicate that 4 D 11 recognises the recombinant proteins containing the HHHHHHGS in ELISA immunoblot, immuno-fluorescence. In addition purified 4 D 11 coupled to beads (affi-gel or CNBR activated sepharose) can be used to purify recombinant proteins under native conditions.

TBC Vaccine

The present invention may be used to provide a recombinant vaccine that is able to induce a humoral as well as a cell mediated helper and cytotoxic immune response against *M. tuberculosis* antigens. In order to elicit a MCH class I restricted cytotoxic immune response, the antigens employed in the vaccine should have access to the cytoplasmic compartment in the presenting cells of the immunised organism. The 85A

protein is one of the member of the antigenic complex 85 ABC and represents the most abundant molecular species secreted *M. tuberculosis* in the culture media and during bacterial infection. Moreover, the protective immune response induced natural infection is mainly directed against the antigen 85A, this protein elicits the production of antibodies and cytokines and stimulate the cytotixic T cells. This would reproduce the processing and presentation pathways occurring during the natural infection of *M. tuberculosis* and the immunisation procedures employing BCG. Whilst not wishing to be bound by theory, it is believed that the development of such a vaccine would require molecular vehicles for translocating bacterial proteins across the cell membrane into the cytoplasmic compartment.

Example 2 - To demonstrate the possibility of using *Antp* as delivery system we have carried out the following experimental activities: (i) cloning of the 85A full length sequence and of a deleted fragment of 60 amino acids (85A.60) in expression plasmids; (ii) development of synthetic genes encoding chimeric proteins in which the homeodomain of Ant has been inserted at the amino terminal end of the proteins 85A and 85.60 (Antp-85A and Antp-85S.60); (iii) expression in *E. coli* and purification of the recombinant proteins; (iv) development of a specific antiserum against 85A; and (v) analysis of the ability of Antp to vehiculate the mycobacterial polypeptides 85A and 85A.60 into the cytoplasm of HeLa cell in culture.

A. Cloning of 85A and 85A.60 coding sequences in the expression plasmid pDS56/RBSII

The full length 85A gene and the sequence of 85A.60 were amplified in PCR experiments using as template DNA extracted *M. tuberculosis* strain H37Rv. The DNA sequence of 85A.60 encodes a polypeptide encompassing the 85A protein from amino acid 100 to 160. This region contains the immunodominant T and B cell epitopes so far identified in the sequence 85A. As primers for the PCR reaction we have used oligonucleotides deduced from the amino acid sequences of 85A and 85A.60. To facilitate cloning in the expression vectors the primer were designed to contain at their 5'

end the restriction sites *Bam* HI and *Sal* I. The amplified sequences were cloned in the expression plasmid pDS56/RBSII thereby generating the plasmids pDS/85A and pDS/85A.60. The amplified sequences 85A and 85A.60 were sequenced to rule out the presence of mutations introduced during the amplification reaction. The expression unit of the plasmid pDS 56 RBSII is under the control of IPTG inducible promotor and yields fusion proteins containing a stretch of six histidines at their amino- or carboxyl-terminal end. The presence of the histidines confer to the proteins a high affinity for nickel ions. The recombinant proteins 85A and 85A.60 were expressed in *E. coli* and purified by affinity chromatography on nickel column. The yields and quality of the proteins recovered at the end of purification process were assessed by analysing the samples eluted from the affinity column with SDS acrylamide gel electrophoresis.

B. Development of synthetic genes encoding the chimeric proteins Antp-85A and Antp-85A.60

The sequence encoding the homeodomain of Antennapedia was amplified in PCR experiments using as template the DNA extracted from embryonic cells of *D. melanogaster*. The homeodomain encompasses the sequence of *Antp* from amino acid 297 to 356 and is encoded by a gene sequence that is not interrupted by introns. As primers we have employed oligonucleotides deduced on the basis of the homeodomain sequences and containing the Bam HI restriction site at their 5' end. The PCR product was sequenced to rule out the presence of mutations introduced by PCR in the sequence of the homeodomain. The restriction site Bam HI allowed the cloning of the PCR product in the correct reading came in front of the sequence of 85A and 85A.60 in the plasmids pDS/85A and pDS/Antp-85A.60 encoded the chimeric proteins Antp-85A and Antp-85A.60 respectively. Both these proteins are characterised by the presence of the homeodomain of *Antp* and the hisidine tail at their amino- and carboxyl-terminal ends respectively. The sequence of *Antp* was introduced at the amino-terminal end to facilitate the correct folding of this domain in the context of chimeric protein sequences. The expression of the chimeric proteins Antp-85A and Antp-85A.60 was assessed by analyising the bacterial lysate on SDS acrylamide gels. In immunoblot experiments the

serum s-85A recognised molecules migrating with the molecular weight predicted for Antp-85A and Antp-85A.60.

C. Functional analysis of the chimeric proteins Antp-85A and Antp-85A.60.

To assess the ability of the chimeric protein to translocate across the cell membrane increasing amounts of bacterial lysates containing Antp-85A and Antp-85A.60 were added to HeLa cells in culture. As control the cells wee incubated with lysates containing the recombinant proteins 85A and 85A.60 that lack the homeodomain of Antp at their amino-terminal end. After 3 hours of incubation at 37°C_2. the cells were fixed in formaldehyde and analysed with the antiserum s-85A to reveal the translocation of 85A epitopes into the cytoplasm. In immunofluorescence experiments the serum s-85A did not react with HeLa cells that were previously incubated with bacterial lysates containing 85A and 85A.60. This finding would indicate that these polypeptides are not able to cross the cell membrane. On the contrary the antiserum s-25 showed a clear cytoplasmic reactivity on HeLa cells that were incubated with the chimeric proteins Antp-85A and Antp85A.60. This last result strongly indicated that the 85A epitopes were localised within the cytoplasm of cells incubated with Antp-85A and Antp-85A.60 thus indicating that the homeodomain of *Antp* has conferred to the chimeric proteins the ability to translocate across the superficial membrane. These findings all together would indicate that the strategy employed to vehiculate antigens into the cytoplasm across the cell membrane can be used to develop an experimental vaccine against tuberculosis. It is anticipated that the purified chimeric proteins Antp-85A and Antp-85A.60 may elicit in the immunised animals both a humoral and cell mediated helper and cytotoxic immune response.

P/cre regulated expression of proteins in eukaryotic cells

The Cre recombinase is a phage derived enzyme that cuts double strand DNA at specific sited (LoxP). When two of these sites are present in opposite orientation the intervening DNA sequence is excised. The Cre recombinase has been shown to work in distantly

related organisms such as bacteria, yease and mammalian cells. *Antp* may be used to deliver the *Cre* recombines inside the cell nucleus and for manipulating genomic DNA at precise locations in a temporal regulated manner. One of the possible applications of Antp-Cre fusion is the development off highly regulated systems for expressing genes in eukaryotic cells. Transformation vectors can be designed for expressing genes under the transcriptional control of a promotor containing, within two LoxP sites, DNA elements functioning as target sequences for transcription repressor factors. In the absence of Cre recombinase the presence of the repressor sequence would not allow the promotor to transcribe the genes cloned in the vector. By adding the Antp-Cre fusion protein to cells the repressor sequence may be excised from the promotor thus allowing transcription to start. This type of expression vector may be extremely useful for studying the function of eukaryotic genes as well as for expressing biologically active molecule in large amounts.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:
(A) NAME: Imperial College of Science Technology and Medicine
(B) STREET: Sherfield Building, Exhibition Road
(C) CITY: London
(E) COUNTRY: UK
(F) POSTAL CODE (ZIP): SW7 2AZ

(ii) TITLE OF INVENTION: Fusion Protein

(iii) NUMBER OF SEQUENCES: 1

(iv) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 60 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

```
Arg Lys Arg Gly Arg Gln Thr Tyr Thr Arg Tyr Gln Thr Leu Glu Leu
1               5                  10                  15

Glu Lys Glu Phe His Phe Asn Arg Tyr Leu Thr Arg Arg Arg Arg Ile
            20                  25                  30

Glu Ile Ala His Ala Leu Cys Leu Thr Glu Arg Gln Ile Lys Ile Trp
        35                  40                  45

Phe Gln Asn Arg Arg Met Lys Trp Lys Lys Glu Asn
    50                  55                  60
```

Claims

1. A fusion protein comprising:

(a) a first region comprising the homeodomain of antennapedia or a variant thereof; and

(b) a second region not naturally associated with the first region comprising a protein of at least 100 amino acids.

2. A fusion protein according to claim 1 wherein the protein is a functional or regulatory protein.

3. A fusion protein according to claim 1 or 2 wherein the protein is an antigen of tuberculosis.

4. A fusion protein according to claim 1 or 2 wherein the protein is an antibody to a malaria antigen.

5. A fusion protein according to claim 1 or 2 wherein the protein is a functional *p53* protein.

6. A nucleic acid encoding the fusion protein of any preceding claim.

7. An expression vector comprising the nucleic acid of claim 6 operably linked to a promoter.

8. A host cell transformed with the expression vector of claim 7.

9. A method for preparing a fusion protein comprising:

(i) culturing the host cell according to claim 8 under conditions which provide for the expression of the fusion protein from the expression vector within the host cell; and

(ii) recovering the fusion protein, which method comprises fusing an amino acid tail to the carboxy terminal end of the fusion protein, which tail is capable of binding to a first substrate and not to a second substrate, and wherein the fusion protein is caused to bind via the tail to the first substrate such that components of the host cell do not bind to the first substrate; and the fusion protein is contacted with the second substrate such that the fusion protein is not bound and remaining components of the host cell are bound to the second substrate.

10. A method for preparing a fusion protein comprising:

(i) culturing a host cell, transformed with an expression vector comprising nucleic acid, operably linked to a promoter, encoding (a) a first region comprising the homeodomain of antennapedia or a mutant thereof; and (b) a second region not naturally associated with the first region comprising an amino acid sequence of interest, under conditions which provide for expression of the fusion protein from the expression vector within the host cell; and

(ii) recovering the fusion protein, which method comprises comprises fusing an amino acid tail to the carboxy terminal end of the fusion protein, which tail is capable of binding to a first substrate and not to a second substrate, and wherein the fusion protein is caused to bind via the tail to the first substrate such that components of the host cell do not bind to the first substrate; and the fusion protein is contacted with the second substrate such that the fusion protein is not bound and remaining components of the host cell are bound to the second substrate.

11. A fusion protein prepared according to the method of claim 9 or claim 10.

12. A pharmaceutical composition comprising the fusion protein of any one of claims 1 to 5 or claim 11 in combination with a pharmaceutically acceptable carrier.

13. A pharmaceutical composition according to claim 12 in the form of a vaccine.

14. Use of the fusion protein of any one of claims 1 to 5 or claim 11 in the preparation of a medicament for the treatment or prevention of cancer, a genetic disease and bacterial or viral infections.

15. Use according to claim 14 for the prevention or treatment of tuberculosis or malaria.

16. Use according to claim 14 for the treatment or prevention of cancer involving a mutated *p53* gene.

17. A fusion protein of any one of claims 1 to 5 or claim 11 for use in an expression system.

ABSTRACT

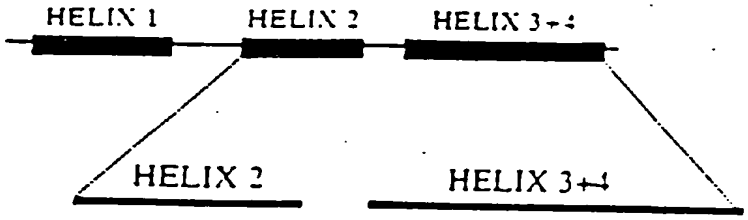
FUSION PROTEIN

A fusion protein comprising:

(a) a first region comprising the homeodomain of antennapedia or a variant thereof; and

(b) a second region not naturally associated with the first region comprising a protein of at least 100 amino acids.

Fig. 1



LTRRRRIEIAHALCLTERQIKIWFQNRRMKWKKEN pAntp

Fig. 2

LTRRRRIEIAHALCLTERQIKIWFQNRRMKWKKEN pAntp

----------Y----------A---------- pAntp 50A

----------PP---------- pAntp 40P2

PCT NO [redacted] / [illegible]

FORM [illegible] 2/9/99

AGENT = [illegible]

THIS PAGE BLANK (USPTO)